Multiplexing biochemical signals

Wiet de Ronde,[1] Filipe Tostevin,[1] and Pieter Rein ten Wolde[1]

[1]*FOM Institute for Atomic and Molecular Physics (AMOLF), Science Park 104, 1098 XG Amsterdam, The Netherlands*

(Dated: August 4, 2010)

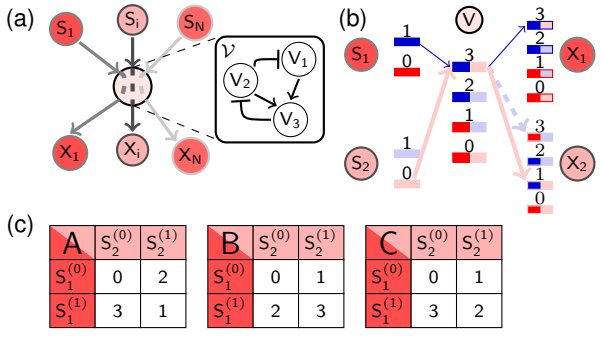
In this paper we show that living cells can multiplex biochemical signals, *i.e.* transmit multiple signals through the same signaling pathway simultaneously, and yet respond to them very specifically. We demonstrate how two binary input signals can be encoded in the concentration of a common signaling protein, which is then decoded such that each of the two output signals provides reliable information about one corresponding input. Under biologically relevant conditions the network can reach the maximum amount of information that can be transmitted, which is 2 bits.

Cells continually have to respond to a myriad of signals. One strategy for transmitting distinct stimuli is to use distinct signal transduction networks . It is, however, increasingly recognized that components are often shared between pathways [1]. Moreover, cells can transmit different signals through one and the same pathway, and yet respond to them specifically. In rat cells, for instance, neuronal growth factor and epidermal growth factor stimuli are transmitted through the same MAPK pathway, yet give rise to different cell fates, differentiation and proliferation respectively [2]. These observations suggest that cells are able to transmit multiple messages through the same signal transduction network, just as many telephone calls can be transmitted via a single wire. Indeed, the intriguing question that arises is whether biochemical networks, like electronic circuits, can *multiplex* signals: can multiple input signals be combined (encoded) *simultaneously* in the dynamics of a common signalling pathway, which are then decoded such that cells can respond specifically to each signal (see Fig. 1)?

The question of how cells can transduce multiple signals via pathways that share components is a key question in biology, since sharing components may lead to unwanted crosstalk between the different signals: from the perspective of one signal, the presence of additional signals constitutes noise. In recent years, several mechanisms for ensuring signaling specificity have been proposed. One is spatial insulation, where the shared components are incorporated into distinct macromolecular complexes on scaffold proteins [1]. Other proposals are based on the temporal dynamics of the system, such as cross-pathway inhibition [3–5] and kinetic insulation [6]. However, these studies only considered scenarios in which the system is stimulated with one signal at the time. Rensing and Ruoff studied what happens when two or three MAPK pathways that share components are stimulated simultaneously [7], but found that one pathway tends to dominate the response, suggesting that multiple messages cannot be transmitted simultaneously. Here we demonstrate that cells can truly multiplex signals: we show that they can transmit at least two signals simultaneously through a common pathway, and yet respond specifically to each of them.

We first have to understand how multiple signals can be encoded in the dynamics of a signaling pathway. Cells employ a number of coding strategies for transducing signals. One is to encode stimuli in the temporal dynamics, such as the duration [2] or frequency [8], of an intracellular signal. In principle, these coding strategies could be used to multiplex signals. Here, we consider what is arguably the simplest and most generic coding strategy cells could choose, namely one in which the signals are encoded in the concentrations of the signaling proteins. We will call this strategy AM multiplexing.

We will consider the biochemical network shown in Fig. 1A. It consists of N input species $\mathrm{S}_1,\ldots,\mathrm{S}_N$ with copy numbers $S_1,\ldots,S_N$, a signal transduction pathway $\mathcal{V}$ consisting of M species $\mathrm{V}_1,\ldots,\mathrm{V}_M$, and N output species $\mathrm{X}_1,\ldots,\mathrm{X}_N$. The copy number of each input species S_i can be in one of K states, $s_i = 0,\ldots,K-1$,



A	$S_2^{(0)}$	$S_2^{(1)}$
$S_1^{(0)}$	0	2
$S_1^{(1)}$	3	1

B	$S_2^{(0)}$	$S_2^{(1)}$
$S_1^{(0)}$	0	1
$S_1^{(1)}$	2	3

C	$S_2^{(0)}$	$S_2^{(1)}$
$S_1^{(0)}$	0	1
$S_1^{(1)}$	3	2

FIG. 1: (a) Biochemical multiplexing: N different signals are encoded in the state of a common pathway $\mathcal{V}$, which is then decoded such that each output species X_i provides reliable information about the corresponding input S_i. (b) Multiplexing is a mapping problem. The states of two inputs S_1 and S_2 are mapped onto the concentration of V, which is then mapped onto states of the output species X_1 and X_2; we require that the two lowest (highest) levels of X_i correspond to the lowest (highest) level of S_i; the dashed arrow denotes a mapping that violates this requirement; levels of V and X_i are colored according to input pattern $\mathbf{s} = (s_1, s_2)$. (c) The 3 unique mappings of $\mathbf{s}$ to v; in panel (b) mapping C is shown.

which are labelled in order of increasing copy number, $S_i^{(0)} < S_i^{(1)} < \cdots < S_i^{(K-1)}$. The input pattern is denoted by the vector $\mathbf{s} = (s_1, \ldots, s_N)$. Similarly, the copy number of each output species X_i can be in one of L states $x_i = 0, \ldots, L-1$ ordered by increasing copy number X_i, and the output pattern is denoted by the vector $\mathbf{x} = (x_0, \ldots, x_N)$. A necessary condition for multiplexing is that the state space of $\mathcal{V}$ is large enough that it is possible to encode the total number of input patterns, K^N, in $\mathcal{V}$.

We imagine that the N input signals are independent, and that the signal transduction network $\mathcal{V}$ replaces N independent signaling pathways. We therefore require that X_i should provide reliable information about the state s_i, but not necessarily about $s_{j\neq i}$; the N different input signals $\mathbf{s}$ simply have to be transduced to $\mathbf{x}$, not necessarily integrated. In general, however, the state x_i will be a function of the states of all the input species: $x_i = f(\mathbf{s})$. This reflects the fact that inevitably there is cross-talk between the different signals because they are transmitted via the same pathway. However, this cross-talk is not detrimental as long as it does not compromise the cell's ability to infer from x_i what s_i was.

Another key point is that while the precise mapping from $\mathbf{s}$ to $\mathbf{x}$ may not be critical for the amount of information transmitted *per se*, this is likely to be important for whether or not this information can be exploited. Let's imagine that the system contains three input species, say three sugars, and that each of these can be in one of only two states, $s_i = 0$ or 1, corresponding to the absence or presence of the sugar; let's further assume that X_i is an enzyme needed to consume sugar S_i. With 8 input patterns X_i can, in the absence of noise, take 8 values, identified as states $x_i = 0, \ldots, 7$. Now, it seems natural to demand that when the sugar S_i is absent ($s_i = 0$), the copy number of enzyme X_i is low, while when S_i is present, the copy number of X_i is high; this means that the four lowest levels of X_i ($x_i = 0, 1, 2, 3$) should correspond to $s_i = 0$, while the four highest levels of X_i should correspond to $s_i = 1$. We therefore require that the mapping from $\mathbf{s}$ to $\mathbf{x}$ is such that the output states $\{x_i\}$ corresponding to input $s_i = j$ are grouped into sets that are *contiguous* and either increase or decrease *monotonically* with j, for each signal i. This leads to a monotonic input-output relation between S_i and X_i for each i. We call this requirement the multiplexing requirement.

In the rest of the manuscript, we make these ideas concrete for a network in steady state with two input species, S_1 and S_2, each of which has either a low ($s_i = 0$) or a high concentration ($s_i = 1$). We take a signaling pathway $\mathcal{V}$ consisting of only one species, V. Multiplexing requires that, in the absence of noise, the four input patterns $\mathbf{s}$ can be mapped onto four distinct states of V, $v = 0, \ldots, 3$, again labelled in order of increasing copy number. These four levels of V lead to four states for each of the two output species X_1 and X_2 (Fig. 1B). As explained above, we require that we can group these four states into two sets, called LOW and HIGH, such that the LOW set, containing $x_i = 0, 1$, corresponds to $s_i = 0$ and the HIGH set, containing $x_i = 2, 3$, corresponds to $s_i = 1$ (or *vice versa*, leading to an inverse input-output relation). To elucidate which mechanisms make it possible to multiplex S_1 and S_2, we note that there exists different ways of mapping $\mathbf{s}$ to v, but, as we will explain shortly, not all of these mappings can be decoded into $\mathbf{x}$ in a manner that satisfies the multiplexing requirement. We therefore first address the question which combinations of mapping from $\mathbf{s}$ to v and decoding from v to $\mathbf{x}$ fulfill the multiplexing requirement, and then we will discuss what encoding mechanisms actually allow for the required mapping from $\mathbf{s}$ to v.

Due to the symmetry in the problem, there are 3 unique ways of mapping the four input patterns $\mathbf{s}$ to v (Fig. 1C). To determine whether there exists a scheme for decoding the signals from v to $\mathbf{x}$ that satisfies the multiplexing requirement, we examine for each mapping all possible network topologies between V, X_1 and X_2, except those that involve autoregulation or mutual repression/activation since these may lead to bistability. In particular, we allow not only for activation and repression of X_1 and X_2 by V, but also for activation and repression of X_2 by X_1, leading to feedforward loops, a common motif in signal transduction pathways and gene networks [9]. In the deterministic mean-field limit the steady-state values of X_1 and X_2 are thus given by

$$X_1 = k_1 f(V; K_\alpha, n_\alpha)/\mu, \quad (1)$$

$$X_2 = k_2 f(V; K_\beta, n_\beta) \times f(X_1; K_\gamma, n_\gamma)/\mu, \quad (2)$$

where k is the maximum activation/production rate, μ is the degradation/deactivation rate, and each regulation function is either an activating or repressing Hill function, $f(V; K, n) = V^n/(V^n + K^n)$ or $f(V; K, n) = K^n/(V^n + K^n)$. The multiplication in Eq. 2 indicates that we assume that at X_2, X_1 and V are integrated according to AND logic [9]. We performed extensive sampling of the space of parameters $k_1, k_2, K_\alpha, n_\alpha, K_\beta, n_\beta, K_\gamma, n_\gamma$ for each of the mappings in Fig. 1C.

Only for mapping C do we find decoding schemes that satisfy the multiplexing requirement [16]. Interestingly, all valid decoding networks are incoherent feedforward loops [9]. Figure 2 illustrates the principle for one such motif. Panel B shows for each of the four input patterns $\mathbf{s}$ the copy number V together with the threshold copy numbers, K_α, K_β and K_γ, while panels C and D show X_1 and X_2 respectively as a function of V. $X_1(V)$ is a simple activation curve with activation threshold K_α. In contrast, $X_2(V)$ starts low and rises around K_β, but then decreases again due to repression by X_1. This non-monotonicity, which is a result of the incoherent character of the feedforward loop, is critical, since this makes

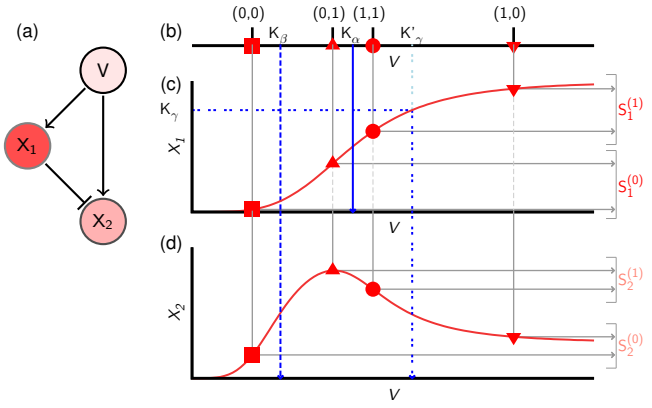


FIG. 2: Decoding V using an incoherent feedforward loop. (a) Network architecture. (b) The values of V corresponding to the four input patterns $\mathbf{s}$ for mapping C (see Fig. 1C), with thresholds $K_\alpha, K_\beta, K_\gamma$ (see Eqs. 1-2). (c) $X_1(V)$; (d) $X_2(V)$. The non-monotonicity of $X_2(V)$ swaps the states corresponding to $(1,1)$ and $(0,1)$ in the mapping from v to x_2.

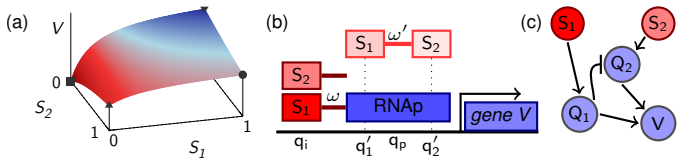


FIG. 3: (a) An input-output relation $V(S_1, S_2)$ consistent with mapping C (Fig. 1C); symbols ■, ▲, ●, ▼ correspond to states in Fig. 2. (b) Architecture of the promoter. (c) A feedforward loop that can generate mapping C.

it possible to swap the order of the states corresponding to $\mathbf{s} = (1,1)$ and $(1,0)$ in the mapping from v to x_2. The key parameters are the activation/repression thresholds K, since they determine where in the state space of V the outputs switch between high and low levels. The precise values of k and n are of less importance, although n should not become so large that $X_1(V)$ becomes Boolean: it is critical that X_1, which needs to be activated by V around K_α to transmit S_1, is not fully activated at K_α: to multiplex S_2, X_1 should reach the threshold K_γ for repressing X_2 only when V has become significantly larger than K_α. Indeed, if X_1 can only take two states, then only three states of V could be decoded, and not the required four. AM multiplexing thus relies on the fact that signals can be encoded over a range of concentrations.

We can now also understand why mappings A and B are difficult to decode: they would require an input-output relation between X_2 and V that rises more than once. This is difficult to achieve in a feedforward loop without mutual repression or activation.

The above analysis shows that it is possible to decode multiple signals simultaneously, provided that the input $\mathbf{s}$ can be encoded in v according to mapping C. The next question is how these mappings, which correspond to particular input-output relations $V(S_1, S_2)$, can be generated. Experiments [10] and modelling [11, 12] have shown that transcriptional regulation can be very sophisticated, allowing for complex logical operations [12]. We indeed find that a simple scheme for transcriptional regulation based on the mechanism of 'regulated recruitment' [11] can generate the required input-output relation $V(S_1, S_2)$, where S_1 and S_2 are now transcription factors that regulate the expression of the protein V. In this scheme, S_1 and S_2 independently activate gene expression by binding next to the core promoter, thus recruiting the RNA polymerase (RNAp), while S_1 and S_2 together repress gene expression by cooperative binding to the core promoter, thereby blocking the binding of RNAp (see Fig. 3). This yields

$$V(S_1, S_2) = \frac{(\beta/\mu)\, q_p(1 + \omega q_1 + \omega q_2)}{1 + q_1' + q_2' + \omega' q_1' q_2' + q_p(1 + \omega q_1 + \omega q_2)}, \tag{3}$$

where β is the maximum expression rate and μ is the degradation rate of V, $q_p = c_p/K_p$ is the concentration of RNAp c_p scaled with its dissociation constant K_p, $q_1 = S_1/K_1$, $q_2 = S_2/K_2$, $q_1' = S_1/K_1'$, and $q_2' = S_2/K_2'$, where K_i and K_i' are the dissociation constants for the binding of S_i to the promoter sites where the RNAp is recruited or blocked, respectively; ω and ω' are factors reflecting cooperative interactions between the respective molecules [11, 12]. We thus conclude that gene regulation networks have the capacity to multiplex signals.

While it is clear that signaling pathways often share common components [1, 2], the logic of signal integration in these pathways has been characterized in much less detail than for gene regulatory networks. It is conceivable that the desired input-output function $V(S_1, S_2)$ could be implemented at the level of a single protein V, using competitive and/or cooperative binding between the three molecules S_1, S_2, V. Alternatively, the required encoding could also be implemented at a higher level of network interactions. For instance, a network in which S_1 and S_2 regulate V via two additional components, Q_1 and Q_2, in an incoherent feedforward loop (Fig. 3C), could achieve the required encoding $V(S_1, S_2)$. In essence, the feedforward loop between Q_1, Q_2 and V can be used to control the ordering of V in the encoding process, just as the feedforward loop between V, X_1 and X_2 can be used to regulate the ordering of X_2 in the decoding step. Since feedforward loops are common motifs in signal transduction pathways [9], we argue that multiplexing can also be implemented in these networks.

The analysis above shows that in principle biochemical networks can multiplex signals in the mean-field, deterministic limit. However, there remains the question of whether signals can be multiplexed reliably in the presence of inevitable biochemical noise. To address this, we

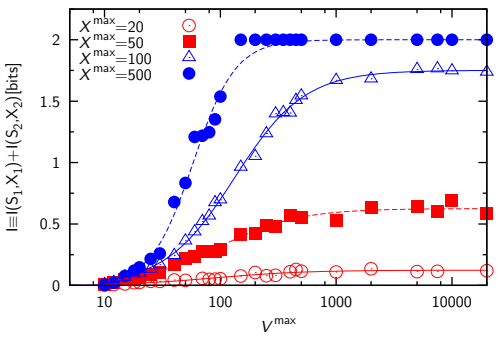


FIG. 4: The transmitted information I as a function of $V^{\max}$ for four different values of $X^{\max}$. I can reach the maximum value of 2 bits provided $V^{\max}$ and $X^{\max}$ are large enough.

estimate a lower bound on the information about two binary signals S_1 and S_2 that are transmitted through the network studied above (Eqs. 1-3). We define the total information $I \equiv I(S_1, X_1) + I(S_2, X_2)$ as the sum of the mutual information for each of the individual signals, $I(S_i, X_i) = \sum_{x_i} \sum_{s_i} p(x_i, s_i) \log[p(x_i, s_i)/p(x_i)p(s_i)]$ [13], where $p(s_i)$ and $p(x_i)$ are respectively the probabilities of S_i being in state s_i and X_i being in state x_i, and $p(x_i, s_i)$ is the joint probability of input s_i and output x_i. Note that in the presence of noise X_i is not limited to 4 states but can in principle take any value. This definition of I makes it straightforward to directly compare the performance of this network with that of two independent pathways. If each of the two input states for each S_i is equally likely then the maximum value of $I(S_i, X_i)$ is 1 bit for each transmitted signal i; the maximum value of I is thus 2 bits.

To maximize the lower bound on I we optimize the network parameters using a simulated-annealing algorithm; we have verified that the final results are robust by varying the initial conditions, and by also using an evolutionary algorithm. We fix the degradation rate of all proteins to be $\mu = 1\text{hr}^{-1}$ and vary n_α, n_β and n_γ between 1 and 4. Values of k_1, k_2 are set such that the maximum mean value of each X_i is $X^{\max}$; similarly, β is set such that the maximum mean value of V is $V^{\max}$. $X^{\max}$ and $V^{\max}$ are varied systematically (see Fig. 4). The threshold parameters K_α, K_β and K_γ are varied over the range $[0, V^{\max}]$ or $[0, X^{\max}]$ as appropriate. We vary q_p, q_i from 10^{-2} to 10^2 and ω, ω' between 1 and 10. For each parameter set we compute $p(x_i, s_i)$ using the linear-noise approximation [14]. Its accuracy was verified by performing Gillespie simulations of the optimized networks [15].

Figure 4 shows that below a threshold copy number $V_c^{\max} \approx 50$ the total information is low regardless of $X^{\max}$ because four distinct states of V cannot be generated. Above $V_c^{\max}$, for large $X^{\max}$ the information I reaches 2 bits, the maximum information about the two signals S_1 and S_2 that could be transmitted via two independent channels. For lower values of $X^{\max}$, I saturates at a value lower than 2 bits, limited by the intrinsic noise in the production and decay of X_i. Importantly, I reaches 2 bits for $V^{\max} \approx X^{\max} \approx 500$, which is well within the range of typical protein copy numbers inside living cells. This shows that biochemical networks can multiplex two signals reliably in the presence of biochemical noise under biologically relevant conditions.

In summary, our results suggest that cells can transmit at least two binary signals through one and the same pathway, and yet respond specifically and reliably to each of them. The proposed mechanism for biochemical multiplexing is based on swapping the order of states during the encoding and decoding steps using incoherent feedforward loops. It is clear that the principle is generic, and could be implemented in signal transduction pathways and gene networks – indeed incoherent feedforward loops are commonly found in these networks [9]. Our predictions could be tested experimentally by simultaneously stimulating two MAPK pathways that share components [1], although perhaps a more controlled experiment would be one using synthetic gene networks. In future work, we will address how more than two input signals can be transduced simultaneously, and how cells can multiplex signals by encoding them into the temporal dynamics of the signaling pathway.

We thank Tom Shimizu for a critical reading of the manuscript. This work is supported by FOM/NWO.

[1] M. A. Schwartz and H. D. Madhani, Annu Rev Genet **38**, 725 (2004).
[2] C. J. Marshall, Cell **80**, 179 (1995).
[3] L. Bardwell *et al.*, Biophys J **92**, 3425 (2007).
[4] M. N. McClean *et al.*, Nat Genet **39**, 409 (2007).
[5] B. Hu, W.-J. Rappel, and H. Levine, Biophys J **96**, 4755 (2009).
[6] M. Behar, H. G. Dohlman, and T. C. Elston, Proc Natl Acad Sci USA **104**, 16146 (2007).
[7] L. Rensing and P. Ruoff, J Theor Biol **257**, 578 (2009).
[8] M. J. Berridge, P. Lipp, and M. D. Bootman, Nat Rev Mol Cell Biol **1**, 11 (2000).
[9] S. S. Shen-Orr *et al.*, Nat Genet **31**, 64 (2002).
[10] S. Kaplan *et al.*, Mol Cell **29**, 786 (2008).
[11] N. E. Buchler, U. Gerland, and T. Hwa, Proc Natl Acad Sci USA **100**, 5136 (2003).
[12] R. Hermsen, S. Tans, and P. R. ten Wolde, PLoS Comput Biol **2**, e164 (2006).
[13] C. E. Shannon, Bell System Tech. J. **27**, 379 (1948); E. Ziv *et al.*, PLoS One **2**, e1077 (2007); P. Mehta *et al.*, Mol Syst Biol **5**, 325 (2009); A. M. Walczak *et al.*, Phys Rev E **81**, 041905 (2010).
[14] N. G. van Kampen, *Stochastic Processes in Physics and Chemistry* (North Holland, Amsterdam, 2001).
[15] D. T. Gillespie, J. Phys. Chem. **81**, 2340 (1977).
[16] Mapping B does allow valid multiplexing solutions, but only with unrealistic parameter values.